

**ABSTRACT OF THE DISCLOSURE**

[0038] A small cache memory can be incorporated with a main memory, such as a flash memory, on an integrated circuit to improve average access times between a processor and the main memory. To minimize cost and complexity, the cache memory may contain only a few words of data. The cache can also allow a suspended transfer with minimal latency when the transfer is resumed. Designing the cache memory to interface with the processor over a standard memory bus permits the cache to be implemented in a system that could otherwise have no cache memory unless the processor and/or memory bus were redesigned.